

# Categorized Web Browsing History

## **BACKGROUND OF THE INVENTION**

### **TECHNICAL FIELD**

The invention relates to Web browsing. More particularly, the invention relates to a method and apparatus for establishing and maintaining a categorized Web browsing history.

### **DESCRIPTION OF THE PRIOR ART**

Browser history is generally shown as one or more flat lists that are column sortable by date last visited, visit count, URL, or page title, *etc.* It is difficult to find specific items in such lists. It is also difficult to find groups of similar sites that were once found and visited. In frustration, most people often resort to starting from scratch and use a search engine to re-find a site in the same way they found it in first place. Browser history is simply unfriendly.

It would be advantageous to provide a simple, user-friendly method and apparatus for establishing and maintaining a categorized Web browsing history.

## **SUMMARY OF THE INVENTION**

The invention provides a simple, user-friendly method and apparatus for establishing and maintaining a categorized Web browsing history. A presently preferred embodiment of the invention uses a directory service, such as the Open Directory

Project, to get category metadata about visited URLs. Metadata thus obtained are used to present a user's personal browser history in a category based hierarchy.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

5

Fig. 1 is a screen display of Web browsing tools that display metadata and allow capturing an image or Web page according to the invention;

Fig. 2 is a screen display of captured Web pages that are stored as part of a media  
10 library according to the invention;

Fig. 3 is a screen display showing Web pages that are part of a browsing history or that have been captured locally according to the invention;

15 Fig. 4 is a screen display showing categories of Web pages acquired from a external source according to the invention;

Fig. 5 is a screen display of a home page for the Open Directory Project; and

20 Fig. 6 is a flow diagram showing a Web browsing history categorization method according to the invention.

25

## **DETAILED DESCRIPTION OF THE INVENTION**

A presently preferred embodiment of the invention provides a simple, user-friendly method and apparatus for establishing and maintaining a categorized Web browsing history. A presently preferred embodiment of the invention uses a directory service, such as the Open Directory Project, to get category metadata about visited URLs. Metadata thus obtained are used to present a user's personal browser history in a category based hierarchy.

Fig. 1 is a screen display of Web browsing tools that display metadata and allow capturing an image or Web page according to the invention. In Fig. 1, a sidebar 11 associated with the Web page cnn.com 12, note the URL 15, shows tools that include "Save Page" 14 and "Stop Watching" 16. A Rating 17, Comments 18, Snapshots 19, and Categories 20 section of metadata 13 is also shown. Those skilled in the art will appreciate that various other types of information may comprise metadata, e.g. links, images, and various statistics.

Fig. 2 is a screen display of captured Web pages that are stored as part of a media library according to the invention. In Fig. 2, the Web page cnn.com 12' is a captured Web page that is displayed from a local source (note the URL 21). The Web page was captured by selecting the "Save Page" button (Fig. 1). In this way, the Web page may be viewed and searched even if the original page is changed or becomes unavailable.

Fig. 3 is a screen display showing Web pages that are part of a browsing history or that have been captured locally according to the invention. In Fig. 3, Web pages that

have either been captured locally or that are part of a browsing history 31 can be viewed in multiple ways, such as a passively personalized Web favorites list Cool Pages 32, or in a time-based history view 33. Note that metadata 13 for a highlighted page are displayed when the page is selected. Double clicking the page  
5 entry in the list retrieves the page from local storage for viewing by the user.

Fig. 4 is a screen display showing categories of Web pages acquired from an external source according to the invention. In Fig. 4, a Categories tab 41 is selected, thereby displaying various categories of pages, *e.g.* News 42, Arts 43, and  
10 Computers 44, acquired from an external source, such as Google. Those skilled in the art will appreciate that any number of categories and sub-categories may be provided.

A particular feature of the invention leverages the existing publicly accessible  
15 databases of URL categorization and uses that information to build a presentation in the browser that users can navigate and explore. For example, the Open Directory Project (ODP; see <http://dmoz.org/>) is a widely distributed database of Web content classified by humans. ODP maintains a database, and exposes a Web interface, that allows navigating through categories to find Web sites (see Fig. 5). This is  
20 similar to the service for which Yahoo! was initially famous. Google leverages the ODP database to allow reverse lookups, *i.e.* finding the category for a given URL. For instance, the Google entry for aol.com shows it to be in the category: Computers > Internet > Access Providers > International > AOL.

25 These categories are structured as trees, where each node has zero or one parents and zero or many children.

As shown in Fig. 6, the key steps performed in the presently preferred embodiment of the invention comprise:

- 5        1. Every time a URL that has not been previously seen is added to history, a reverse lookup is performed to find the chain of categories for the URL (100).
- 10       2. For some URLs there is no category in the ODP database. For instance, the URL might represent some obscure page in a Web pod. In that case, the system iteratively retries the reverse lookup using a less specific part of the URL until a category is found (110). For instance, if no category is found for <http://www.cnn.com/stories/entertainment/12345.html>, then the system searches for a category for <http://www.cnn.com/stories/entertainment> and, if that fails, then <http://www.cnn.com/stories>, and so on.
- 15       3. Standard relational database technology is used to store and query a local representation of categories (120). A table row is used for each category node. Each row has a parent field to represent the node to the left in the category chain. For instance, given a chain "Computers > Internet > Access Providers > International > AOL" the 'AOL' node has the 'International' node set as its parent. In another database table, the relationship between each URL and its one specific category node is stored, although a URL may be mapped to more than one category as appropriate or as determined by a user.
- 20       4. Standard relational database technology is used to store and query a local representation of categories (120). A table row is used for each category node. Each row has a parent field to represent the node to the left in the category chain. For instance, given a chain "Computers > Internet > Access Providers > International > AOL" the 'AOL' node has the 'International' node set as its parent. In another database table, the relationship between each URL and its one specific category node is stored, although a URL may be mapped to more than one category as appropriate or as determined by a user.
- 25       5. Standard relational database technology is used to store and query a local representation of categories (120). A table row is used for each category node. Each row has a parent field to represent the node to the left in the category chain. For instance, given a chain "Computers > Internet > Access Providers > International > AOL" the 'AOL' node has the 'International' node set as its parent. In another database table, the relationship between each URL and its one specific category node is stored, although a URL may be mapped to more than one category as appropriate or as determined by a user.

4. Given such a category path, add rows for any nodes that are not already in the local database, and then set the parent links appropriately (130).

5. Using standard relational database querying, find the children categories of any given category by querying for all the category nodes whose parent is the given category (140). Similarly, find all the URLs for a given category by querying for all the URLs associated with that category (150).

This data model permits one to build interactive user interfaces that allow users to navigate among the categories corresponding to URLs that have been previously visited while browsing. Unlike the ODP user interface, it does not include categories that correspond to pages the user has never visited.

The user interface for this hierarchy of data can be displayed as a tree list, as with the Windows file system Explorer, or as a series of Web-style pages, as with the OPD Website. The resultant user interface allows users to re-find Web sites by drilling down through familiar categories. It also allows users to discover groups of previously visited pages which happen to fall into the same category. This grouping can be very useful to users. Optionally, from any given category display the user interface can supply a link back to the full ODP Web page for that category. Thus, within the same user interface, the user can focus on previously visited pages and then navigate to a wider view of similar pages that might be worth exploring.

Although the invention is described herein with reference to the preferred embodiment, one skilled in the art will readily appreciate that other applications may be substituted for those set forth herein without departing from the spirit and scope of

the present invention. Accordingly, the invention should only be limited by the Claims included below.